CHAPTER 2: LITERATURE REVIEW
2.1 Introduction

This chapter describes and discuss the technique used in producing a timetable especially those at university level. The actual problem faced by researchers in this field is the lack of research working papers on university timetable that are published. Most of the timetable system produced can only be used at the university involved in the research, as each university has its own needs and requirements that differs specifically.

2.1.1 University Timetable

University timetable ensures that each class has its subject and the room to be used. The actual content of the timetable in part is stipulated by the respective faculties. The total number of time a particular subject is taught in a week is normally identified earlier. Each class has a group of students, who must be present when its time for studies and proceeded till the end of the study time. A lecturer is responsible for each subject.

Souza (2000) defines a timetable as the total schedule of a specific teaching, which will be attended by a group of students and the lecturer, at a specified time. It requires specific resources such as, room, teaching aid and the like. Scheduling is parallel to the resources made available besides fulfilling other needs.
2.1.2 Formal Definition on Timetable

The timetable involves a set of incident \( E = \{ e_1, e_2, \ldots, e_v \} \) as an example: examination, seminar, project, meeting a set of time to the said incident \( T = \{ t_1, t_2, \ldots, t_s \} \)

The maker of the timetable also need to know a set of place where the incident occurred \( P = \{ p_1, p_2, \ldots, p_m \} \) and a set of agents in conduct of the incident \( A = \{ a_1, a_2, \ldots, a_n \} \) for example: lecturer, tutor.

Allotment is made based on four verses \((e, t, p, a)\) whereby \( e \in E, t \in T, p \in P \) and \( a \in A \).

In brief, it means: the event \( e \) occurs during \( t \) time at \( p \) involving agent \( a \).

The formal definition for lecturer timetable as given by Carter et al (1995) is as follows:

A set of course is given as \( K_1, \ldots, K_q \); the course \( K_i \) has \( K_i \) number of teachings each time. The total time is \( p \). The students are divided into group \( r \) namely \( S_1, \ldots, S_r \) whereby to each \( S_1 \) all students have opted for the same course \( l_k \) is the maximum number of teachings that may be scheduled at the time known as \( k \) (according to the number of rooms available)

Then make \( y_{ik} = 1 \) if the teaching of course \( k \) is scheduled or the time \( K_i \) and \( y_{ik} = \emptyset \) if otherwise. Maximum

\[
\sum_{i=1}^{q} \sum_{k=1}^{p} K_i y_{ik}
\]

Subject to the following constraints:

\[
\sum_{i=1}^{q} y_{ik} = k_i \quad (i=1, 2, \ldots, q)
\]

\[
i=1
\]
p
\[ \sum_{i=1}^{p} y_{ik} \leq k_i \quad (k=1,2,\ldots,p) \]

p
\[ \sum_{i \in S_i} y_{ik} \leq 1 \quad (l=1,2,\ldots,r, \quad k=1,2,\ldots,p) \]

Costs $Y_{ik}$ in objective function (1) is the measurement to $= 1$; a high value may be imposed if the teaching of course $k_i$, occurs at $k$, or 0 if otherwise.

The formal definition for university timetable is given in detailed by Carter et al (1995). In event that the following resources and time slot exist as follows:

- A set of lecturers \{\textit{t}_1, \textit{t}_2, \ldots, \textit{t}_n\}
- A set of subjects \{\textit{c}_1, \textit{c}_2, \ldots, \textit{c}_m\}
- A set of classes \{\textit{r}_1, \textit{r}_2, \ldots, \textit{r}_q\}
- A set of time \{\textit{p}_1, \textit{p}_2, \ldots, \textit{p}_s\}

As a preliminary step, construct a three verse set <\textit{t}_i, \textit{c}_j, \textit{r}_k> to represent the requirement which are (\textit{i}=1, 2, \ldots, \textit{n}), (\textit{j}=1, 2, \ldots, \textit{m}) and (\textit{k}=1, 2, \ldots, \textit{q}). Each set indicates lecturer \textit{t}_i lecturing the subject \textit{c}_j in class \textit{r}_k. Further, determine each three verse set to time \textit{p}_1 which is (\textit{l}=1, 2, \ldots, \textit{s}). This will form a four verse set <\textit{t}_i, \textit{c}_j, \textit{r}_k, \textit{p}_1> which represents the stipulation that lecturer \textit{t}, lectures the subject \textit{c}, class \textit{r} and at \textit{p} time.
For every two four verse set, for example \(<t_e, c_t, r_g, p_h>\) and \(<t_w, c_x, r_y, p_z>\) the constriction below must be fulfilled:

\[
\begin{align*}
\text{If } t_e &= t_w \text{ hence } p_h \neq p_z & (1.1) \\
\text{If } c_t &= c_x \text{ hence } p_h \neq p_z & (1.2) \\
\text{If } r_g &= r_y \text{ hence } p_h \neq p_z & (1.3)
\end{align*}
\]

Constriction (1.1) ensures that no lecturer will schedule more than once in each period. Constriction (1.2) ensures that no subject will be scheduled for the same time and constriction (1.3) ensures that no class is scheduled for more than once for each period.

### 2.2 Heuristic approach

The heuristic approach is one of the first automation approach used in scheduling problems. This approach attempts to obtain the best solution selection despite there being no guarantee to that effect. Looking for the best solution will take a considerable time as problems pertaining to scheduling has many probability which must be evaluated. Through this approach, the number of probability may be limited so that solution may be made within an appropriate time.

Souza et al (2004), attempts to arrange classes suitable to the lecturer that no lecturer will lecture in two classes at the same time slot. This approach may solve several constrictions such as subject that have various time, joint subject or carried out at one time, the distribution of subject in a week and provision of rooms required.
Al-Khwarizmi commences by selecting the most difficult subject to be generated. This difficulty is overcome by using a function that calculates the number of time slot available and the number of time slot needed. Computation is done by using both these parameter that is the number of time slot available is subtracted by the time slot needed.

If the difference is positive in nature, hence several time slot selection will be available, but if the difference is zero, it indicates the time slot available is only sufficient to be given to the subject.

The allotment of subject to time slot is done with the least interference to other subjects. This interference are defined as reduction of time slot available and the number of time slot needed for all subject that are directly affected from the allotment of the said subject.

In event that there are subjects that could not be given time slot, the process will be ended. Hence, the system does not guarantee that all subjects will be given time slot. Besides that, low constrictions are not taken into account, as teaching time slot according to the priority of lecturer and the allotment of subjects for the week does not indicate a good distribution.

Bresina (1996) was among the early researchers who use this approach. He made several modification in the manual approach conducted at universities. It commences by allotting the combination of lecturer – subject without the existence of any conflict between the resources available. In this approach, the search process will allot difficult
subjects first. In event that there are subjects that are not allotted, a change process between the subjects will be done. This is done by shifting the subjects that are already allotted to a time slot that may be suitable to enable the said new subject to be allotted a time slot.

Schaerf (1999) provides several steps on how the heuristic approach is conducted. Firstly, the most difficult subject will be selected first. There are several criteria in ascertaining the difficulties of a certain subject. Among them are:

- Subjects with time slots that are more than one period
- Subjects are carried out simultaneously or incorporated with several other classes
- Subjects that need to be pre-allotted. For example, if the said subject only has one time slot selection in a week.

Secondly, the selection of time slot for each subject. Selection is based on the time slot that is less critical, that is not in anyway depriving the opportunity of other subjects in event that the said time slot is selected. As an example, a subject with a selection of two time slot, in event that the first choice or selection is made, the selection for other subjects will be affected and if the second selection is made, it will not give any effect to other subjects. Hence the second selection is less critical than the first choice.

The disadvantage of using heuristic approach is that there are subjects that could not be allotted with time slots. To solve this problem several actions will be taken. They are:
• The process will be terminated

• Subjects without time slot will be given time slot and the process will generate the following subject.

The process will return to the allotment of subjects prior to that.

2.3 Integer Programming Approach

Linear programming is a mathematical technique to aid in obtaining the most effective decision by utilizing all available resources. In general definition, it is also known as mathematical programming. Integer programming yield solutions for linear programming hitches. Integer programming faces the same setback as linear programming by reason of nearly every variable must be optimize with positive integer.

For example, in the complication of linear programming, the variable can use fraction numbers (not round numbers). However most of the difficulties in the scheduling issue uses variable with integer value. Another example is a classroom must comprise of a single lecturer and not a fraction of lecturers. The definition of linear programming:

1. Maximizing (or minimizing) one linear function to a definite variable. The said functionality either maximizing or minimizing is known as objective function.

2. The definite variable value must adhere to existing constraints. Each constraint must be either linear equation or linear irregularity.

3. All variable must not have negative value. Negative value for physical quantitative such as lecturers, classrooms and subjects is not permitted.
Integer programming undertook the attempt to solve scheduling issues. Dahal et al (1997) defined that problem must be divided into two portions which is subject collection against time followed by unification of subject and classroom.

Difficulties faced when using this approach is the event of too many or all variables in the formula is integer in value and normally the size of the problem is big. It involves extensive calculation time and probability to solve the situation without utilizing mathematical approach. Carter (1995) commented that each problem must be solved mathematically, the number of variables and constraints will be too complicated to manage when facing large problem.

Research was conducted to lessen the amount of variables before applying linear programming. Souza et al (2000) presented two papers involving examination schedule by lessening the technique. It involves 400 subjects and 30 hours a week. By utilizing the mathematical formula, a total of 12000 variables excluding constraint generated. He reduces the size of the problem by creating conflicts between subjects in the attempt to minimize the amount of variables.

Students were gathered according into stream as basis to form subject group. The classroom is gathered according to size and capacity. This will form a new problem which is similar to common problem which the solution will be simpler but by lessening the search scope, the probability would not generate good or feasible results.
According to Souza, this collection will reduce the size of the problem but yet too big to be solved by using integer programming. Therefore, researchers use the method of Lagrange and Relaxation which is dissimilar constraint is included in the objective function by using Lagrange multiplications.

2.4 **Graph Coloring Approach**

Generally most timetable software is based on graph delegation. Graph can be defined as dots connected by lines. In a particular problem, dots on the graph will represent events and lines will correlate with those events. Two adjacent dots (contains conflict) if linked with the same line (common edge). For example, Physics and Chemistry linked with a single line due to both subject cannot be accommodated at the same time.

The graph can be separated by coloring the dots, where no adjacent dots (contains conflict) have the same color. At the end of the coloring process, the total number of colors used demonstrates the separation that exists. This process is known as chromatic counts, a technical term to show the minimal colorization in graph coloring.

This separation is equivalent with the total time occupied to schedule one set of subject, which either one conflicted or inherits the same time schedule. Therefore, the subject must be joined without conflict even though the schedule is not feasible.

The difficulty of creating the timetable can be represented using graph. The subject will be represented by points or dots with constraints with other subjects (e.g. two subjects
can not be scheduled at the same time) are represented by the line. Each separation inside the graph will consist of subjects that can be scheduled at a similar timeslot. The total number of chromatic on graph shows the minimal slot time needed to reschedule all subjects.

Figure 2.1 represents a simple model in developing timetable utilizing the graph coloring approach. Each subject is drawn as a single point and each pairing subject can not be scheduled at the same time. It will be represented by a line splitting between the two points. Furthermore, each point (subject) is colored particularly to show no pairing points connected by a single line that shares the same color.

All colors represents time differences, therefore no conflicting subject will be scheduled in pair. Subject physics is colored differently as compared to mathematics, science, geography and even chemistry. This bears the meaning that physics will be allocated a different time compared to other subjects.
Nevertheless, difficulties still coexist with the usage of this graphical model. It seems hard to represent certain elements such as connecting two different subjects with different time (Time differences of each subject). The relationship is limited only to the subject that can not be scheduled together or vice versa. However the situation can be uplifted by using a weight age as tested by Hertz et al (1987).

2.5 Network Streaming Approach

This method tries to allocate time for the combination of lecturer-classroom whereby no conflicts generated in the schedule. This method solves one problem at a time. At each time, one stream of problem is solved. It is only one node from a source is connected by
the line for other remaining nodes. The classroom nodes are connected to suitable lecturer
nodes which are then connected to other lecture nodes underneath it.

During each steps, priority must be given to the combination of lecturer-classroom
moreover is the next problem to solve. This priority is specified through the functional
cost, which is connected by the lines. Therefore, the cost flow can be minimized.

In real life scenario, certain requirements such as multiple time subjects, subjects with
certain classrooms and concurrent classes are pre-allocated. These combinations will be
extracted and the sources (lecturers or classrooms) involved will become useless for the
next schedule generation.

The researcher is of the opinion that those cases are extremely rare and would occupy
less then 5% from the total combination and would not inflict difficulties during schedule
generation. In tests conducted, the situation above involves a small portion of difficulties
and the combination also cannot be determined. According to the analyst, the
combination is allocated in the timeslot manually which involves a fraction of the
timetable.
2.6 Logical Constraints Arithmetic Approach

Logical constraints arithmetic is a technique that inherits some similarities with Al-Khawarizmi’s Heuristic. Waterloo PROLOG (WPROLOG) is used by some universities scheduling system without having any conflict. Logical programming was used to represent the constraints while pattern matching and backtracking was used for the search engine.

The timetable was divided into two sub problems. First, the scheduling for the lecturer-classroom-subject is given a time slot and secondly the classroom scheduling will be joining the previous state. The use of logical constraints arithmetic could ensure that all high constriction condition was fulfilled to ensure that the timetable is feasible.

Concentrating on the classroom scheduling, the rooms that suits the needs for any configuration, faces some possibilities of not being available. However, bigger classroom may have the probability of vacancy. Classrooms are classified by size and are given a running number. Those subjects that are unsuccessful with any allocated rooms can be reconsidered to loosen the size constraints by allocating bigger room size with bigger running number. A mechanism known as size matching will determine the maximum room size for any course that may be requested. This procedure was successfully implemented in a university in Ottawa.
2.7 Knowledge Base Approach

Davis (1991) stated that the most advance search engine is rule-based. From his opinion, the best solution for the future is the development of intelligent system starting with the recreation of timetable system with several adjustments for a more feasible system.

Carter et al (1995) sparked a new general paradigm for the solution of resource allocation by applying a smart system. The paradigm has four main components:

1. Redefinition of all resource allocation difficulties. Among them are:
   - Activities or jobs must be executed
   - All resources needs to execute activities
   - Priority – Straight allocation for activities and resources
   - Allocation – Joining sources to activities
   - Constraints – Limiting allocation

2. Rules for allocation procedures are as follows: -
   - Adherence to strict rules (high constraints). Other recommended rules (low constraints) to be fulfilled as possible.

3. Strategies for sources allocation are: -
   - Advance allocation is the most difficult execution for joining or allocation.

4. Ensuring all local variables, if fails during execution, must be retraced to the previous activities to cancel and allocate it to another current activity.

On the other hand, De Werra (1995) employed two-execution phase in the attempt to minimize searching difficulties. Problems are broken into sub problem of weekly and
daily. In the first phase, subjects will be predetermined for its days and the second phase will allocate time for the selected days. If the situation cannot be resolved, the first phase will be repeated and a different day will be selected. Priority is given to a time and selection of time slotting activity for allocating a reasonable time slot.

Knowledge base defines that every activities and also time slots are represented separately. Every time slots has a knowledge base. It will be revoked during joining or allocation process or a process to resolve any issues. This knowledge base will determine whether the time slot can be used or not. The knowledge base is divided into several resolution constraints mechanism.

Knowledge base manipulation is mentioned through the usage of subject shifting and subject exclusion. Exclusion is a mechanism to broaden the searching scope utilized in the paradigm.

Resolution constraints mechanism is used to show how the time slot can be implemented to activities that yet to be combined. If the time slot cannot be used for any activities, others that share the same resources are extracted to enable the first activity to be joined. This is for the purpose of shortening the search extent due to the previous joining contains limited relationship between the activities. The previously extracted activity will be allocated to another time slot.
2.8 Tabu Searching

Tabu searching was introduced by White (2000) towards obtaining optimal universal solution for any complication. Minghe (2006) also utilized this technique to obtain high optimal solution for combination optimization difficulties. This repeated approach is in the direction of getting minimal functional objective.

The generation of early feasible timetable that is followed by generating other timetable is known as neighbor. It will sustain small changes from the timetable. Later on, it will be reshuffled to allocate the best solution for the neighbors. To prevent movement for the previous allocated time slot, a certain mechanism is applied to prevent movement for the first arrangement at the end of the movement. However, if the movement is still giving the best optimal result, therefore there will be a rollback to the previous solution.

All the forbidden solution is known as Tabu searching. Any movement in relation to the solution is called Tabu movement. Any Tabu status inherited by which ever solution can be eliminated if the solution is assuming that the solution is obtaining better improvement. This repetition will stop when achievement to the closest functional objective or total repetition have reached the maximum number.

Inside the report by Minghe, he succeed to generate the timetable without facing any conflict unless it involves one or many lectures and students that has more than one concurrent subjects moving between two classes. Functional objective will consider all conflicts in the timetable.
Timetable or scheduling is not generated randomly but on the other hand is done continuously by joining. Therefore lesser conflict generated. This is to enable the neighboring schedule to repair its functional objective. This process basically improves the quality of the early scheduling by minimizing the amount of conflicts. What was done by Sun involves 288 courses for one session (one subject for a week), 143 lectures, 67 classrooms and 1792 students. The analysis had insufficient information to be commenting that the attempt was successful.

2.9 Annealing Simulation

Annealing simulation was suggested by David et al (1991), whereby it was a local search technique that minimizes the probability to transpire universal concentration.

Annealing is a process of heating physical solid objects and gradual cooling. The gradual cooling process ensures all atoms in the solid object to be assigning a suitable spot. At high temperature, atomic particle is unbounded. If it cools relatively fast, the atoms cannot move and obtain unsuitable location. This involves low resolution compared to gradual cooling (annealed).

At any given temperature, the structure of the new atom is accepted if the energy of the system is still low. If the energy is high, the structure might be reconsidered if the increase probability is low compared to the expected temperature given. In the timetable, Abramson (1991) substitute atom with the subject element and energy system with timetable costs. The cost for the timetable is predetermined through functional objective
that measures the schedules infeasibility. Cost calculation is based on clashes of subjects, lectures and rooms. The temperature and cooling rate is predetermined. The temperature will control possibilities of accepting higher cost increase. The cost adjustment is calculated with two components:

1. Cost for removing subject from time.
2. Cost for inserting subject into time.

The alteration in cost is the differences of both components. Subject will move into a new time if the cost alteration is accepted and also if the decrease of increase of cost is accepted at the particular temperature. The cost of rearrangement of the subject involves costs from classes, lecturers and rooms. The cost also applies to the reinsertion of subject that involves classes, lecturers and rooms. It the changes of reinsertion of subject into time, the number of classes is greater than 0, therefore the cost for classes is 1. After relocation of lecturers, if more than 1 lecturer exists, the lecture cost is 1. Room cost and cost for subject reinsertion is based on the technique. This way, the determination of changes is the cost that is excluded from the overall cost of scheduling.

After several repetitions, temperature will decrease and this process will repeat. The overall process will end based on two conditions. Firstly, if the cost for scheduling is 0 (represent no collision), the system will stop as a result of a feasible timetable. Secondly, is when the scheduling cost is constant in several repetitions.

Abramson conducted several attempt using nine sets of timetable which was generated randomly with real time data from university. There was no special need in the first nine
sets of data rather the tenth data required actual requirements of a university. Execution time became bigger as the problem increases. The nine sets of data include 20 lectures, 20 classes and 600 requirements that took 5548 seconds to execute and real time data from the university took 14 hours. Most of the data sets are executed several times for better results.

2.10 Al-Khawarizmi Genetic

2.10.1 The History of Al-Khwarizmi Genetic

John Holland is an early researcher in Al-Khwarizmi Genetic. He published a book known as “Adaptation in Natural and Artificial Systems” and which was followed by working papers and thesis published by other researchers. Two seminars in Al-Khwarizmi Genetic: “Proceeding of the International Conference on Genetic Algorithms and their Applications” are being held every two years since 1985 and a theory seminar on “Foundation of Genetic Algorithms and Classifier Systems” was held annually since 1990.

There are several journals publishing topics on early research done on Al-Khwarizmi Genetic and Al-Khwarizmi evolution such as “Adaptive Behaviour and Evolutionary Computation” (both are from MIT Press), the journal “Complex System” contains article mostly on the Al-Khwarizmi Genetic background and “the Annals of Mathematics and Artificial Intelligence” which was specially published on Al-Khwarizmi Genetic and research in relation to it. Electronically, Alkhwarizmi Genetic has a list of academic
2.10.2 Definition of Al-Khwarizmi Genetic

Gardon (1995) defines Al-Khwarizmi Genetic as a problem solving method using the principles of natural population genetic. The structure on population knowledge representing several solutions will be maintained. Subsequently allowing the population to multiply by way of competition and control variation.

Al-Khwarizmi Genetic differs from Al-Khwarizmi Evolution based to two general characteristics:

- Al-Khwarizmi Genetic operates in the search space using the genetic operator in fixed-length string. Al-Khwarizmi Genetic does not directly use the problem space on the other hand information is coded in the problem space.
- Al-Khwarizmi Genetic uses the unique re-amalgamation operator in Al-Khwarizmi Genetic which is crossover in fixed-length string.

The said difference is based on the definition of Al-Khwarizmi Genetic which was given by Goldberg (1989). He states that Al-Khwarizmi Genetic is a search on Al-Khwarizmi which is based on the selection mechanism and natural genetic. It includes the process that `the most active will prevail’ between the sequence structure and the information structure which changes at random to form the Al-Khwarizmi search which has a slight similarity with the natural process in the development of the human generation.

He further states that, in each generation, an artificial sequence is created using bytes from the older generation, this new generation will attempt to form its best. Although at
random, Al-Khwarizmi Genetic is not a simple random. On the contrary, it will fully use the initial information to speculate against the new search point with the purpose of producing better performance.

How is Al-Khwarizmi Genetic different from other search methods and optimization? Goldberg (1989) has the following views.

- Al-Khwarizmi Genetic uses a coded parameter set.
- Al-Khwarizmi Genetic evaluates objective functions.
- Al-Khwarizmi Genetic by using the population set.
- Al-Khwarizmi Genetic makes selection using probability and does not use the rules as specified.

From the above four characteristics, huge problems may be settled using the Al-Khwarizmi Genetic as problem parameter will be coded as a fixed length sequence. Al-Khwarizmi Genetic operates directly using only the objective functions without specific additional information.

In addition, the use of `database’ or population will reduce the incidence of being trapped in local optimum. Finally, by using probability as a method of selection, it will actually assist Al-Khwarizmi in moving freely in the search space, with the intention of obtaining better development. Al-Khwarizmi Genetic may be moved several times using the same parameter, to possibly obtained the best solution.
2.10.3 Population and Generation

Individual population in Al-Khwarizmi Genetic are always of fixed size which differs from natural population. The initial population is normally generated in a random manner, even though there is a possibility where an individual is coded higher than average. Besides that, Al-Khwarizmi Genetic experience generational – a new population will be formed from the old population through re-production, crossover and mutation, and the old generation will be removed.

Figure 2.2 show a flow chart on the main function in Simple Genetic Algorithm (SGA) introduced by Goldberg (1989)
The population size or number of time table generated relies on the computer system. Figure 2.3 indicates how the operator in SGA functions. For example, if 10 initial population is generated, each population will be evaluated in terms of its fitness. This 10 families will then produce 10 new generation children and the old generation will be removed (this most certainly does not happen in the human race) This process will

**Figure 2.2: Simple Generic Algorithm (SGA)**
continue for a period of 100 generations. This entire procedure will repeat if a particular run is done 10 times, and select the best timetable from each run. The birth of a new generation will be shown by the use of genetic operatic.

![Flowchart](image)

**Figure 2.3: Birth of New Generation**

The relationship between generation may be shown through Figure 2.4. SGA uses population that does not clash between one and the other. Every new generation is created entirely from the earlier population, and later the old population will be disposed. This differs from the natural human population, where is there is a possibility of a new generation having a high number in total population. On the contrary in Al-Khwarizmi
Genetic, the population size is fixed, each generation consists of the same number of individual as before.

![Figure 2.4: Relationship Between Old and New Generation](image)

### 2.10.4 Selection and Reproduction

The most simple rule used in family selection is by the roulette wheel selection. For example, table 2.1 shows population with four sequence, each has a length of 5 bytes. The sequence codes the integer \( x \), and the objective function (fitness) is \( f(x)=x^2 \).

#### Table 2.1: Objection Function

<table>
<thead>
<tr>
<th>No.</th>
<th>Sequence</th>
<th>Fitness</th>
<th>Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01101</td>
<td>169</td>
<td>14.4</td>
</tr>
<tr>
<td>2</td>
<td>11000</td>
<td>576</td>
<td>49.2</td>
</tr>
<tr>
<td>3</td>
<td>01000</td>
<td>64</td>
<td>5.5</td>
</tr>
<tr>
<td>4</td>
<td>10011</td>
<td>361</td>
<td>30.9</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1170</td>
<td>100.0</td>
</tr>
</tbody>
</table>
The divisions in the wheel (Figure 2.5) represent each individual in population. As an example, each sequence has 14.4% division in the wheel upon division of total fitness (169) which is divided by the entire total fitness (1170). The selection of individual is based on each vacancy in the new population. If the new generation population is still incomplete, the wheel will spin to select new individuals.

The ball will stop when a number at random between 0.0 and 1.0 represent the circumference of the will is selected. This random number will be compared to the individual divisions in the wheel. If the individual division exceeds the random number, then the said individual will be selected. Upon the selection of an individual, this does not mean that it will not be selected again. The wheel will remain as the original and will spin again until all new population is filled up.
2.10.5 Chromosome

The use of a fixed length sequence in Al-Khwarizmi Genetic are influenced by the individual’s biology organism genetic heritage coding in the DNA which forms the chromosome. In Al-Khwarizmi Genetic, each individual in the population consist of a fixed length sequence (also known as chromosome), normally using the multiple coding (string). The possibility sequence may also be represented as an integer or decimal (decimal-coded) and possibility also in the matric form.

\[
\{ 0,1,1,0,1,0,0,1,0,1,\ldots \} \quad \text{Multiple Coding}
\]

\[
\{ 2,7,-1,0,0,11,20,6,16,\ldots \} \quad \text{Integer Coding}
\]

\[
\{ 40.77,-11.0, 0.027, 3.9, 2.77,\ldots \} \quad \text{Decimal Coding}
\]

2.10.6 Fitness

The fitness of biological organism is measured from its relationship with the environment. Fitness ascertains the span of its survival, its opportunity in reproduction and finally the number of offspring available. In Al-Khwarizmi Genetic the individual’s fitness is ascertain by decoding of the chromosomes in the problem space. The value of objective function found as a result of possible changes done, is known as the individual’s fitness. The value of this fitness is used to ascertain the extent on individual that is allowed to reproduce to the future generation.
2.10.7 Crossover

What is the population evolution from one generation to another that follows? As with the result from reproduction as a function to fitness, many biological organism reproduce sexually to allow genetic characteristics of two organism to influence the offspring that may be produced, that is, the possibility of an offspring having a better level of fitness then its family.

In Al-Khwarizmi Genetic, the family selected for reproduction process (with a specific probability), not only duplicating the chromosome to that of their children but both families will unit the chromosome using the crossover operator. Hence their child’s chromosome is the result of the unity of their families’ chromosome.

As an example, Figure 2.6 shows that if the sequence has a length of k, hence the crossover point j is selected between 1 and k-1 (crossover cannot be done at 0 or k). All byte between 0 and j-1 will remain but between j and k-1 will be swapped between both the sequence.

```
\[
j - 1 \quad j \quad k - 1
\]
```

![Figure 2.6: Crossover](image)

Figure 2.6: Crossover
Two sequences from the above example, in event that:

\[
\begin{array}{c|c}
0 & 1 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
\end{array}
\]

Namely \( k = 6 \) and \( j = 5 \) (shown through lines), this crossover will produce:

\[
\begin{array}{c|c}
0 & 1 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 \\
\end{array}
\]

### 2.10.8 Mutation

In the real world, mutation is a different factor which influences the genetic heritage of biological chromosome. Most mutation process will result in a vast change to organism, but occasionally one of it will be beneficial, that is an increase in organism fitness will happen, and this results in, the production of genetic from the said mutation will be carried by many organism when the generation increases.

There is a possibility in the Al-Khwarizmi Genetic that mutation occurs to the offspring chromosome during reproduction with random changes (at low probability) of each byte (or non-multiple character) of the said chromosome. Meanwhile, probability of mutation occurring at a low rate is due to the operator which is not as important as the rate itself – fitness due to reproduction dan crossover. For example, upon the occurrence of crossover, the possibility of a new sequence is:

\[
1 & 1 & 0 & 0 & 0 & 1
\]
In the event that mutation probability is 0.001. For the first byte, a function will generate a number at random between 0 and 1. If the said number is lesser that the mutation rate (0.001), hence in the first byte mutataion will occur. The same procedure will be repeated in other bytes. If only the first byte experience mutation, hence the new sequence is:

0 1 0 0 0 1

Al-Khwarizmi is robust and general in nature, that is, it does not use specific problem information – Al-Khwarizmi focuses on development of Al-Khwarizmi Genetic for all problems. It is also known as traditional Al-Khwarizmi Genetic.

2.10.9 Al-Khwarizmi Genetic in Timetable

Abramson (1991) implemented this technique for solving problems in timetabling at university. Cost is accounted from the total collision in the timetable and if cost is empty (no collision), therefore the timetable is accepted. The cost for time includes subjects, lectures and classrooms are given different weight age suitable for its importance. For instance, optimization problem is to minimize the cost by lessening the amount of collision occuring in the timetable.

Timetable is represented with a set of diagraph (a combination of lecturer-subject-classroom) in which will be joined with time to create a complete diagraph. The genetic definition described a continuation of time represents chromosome and diagraph represents gen. Crossover executes method either having one or two crossovers points or by having a uniform crossover where the variable time in the diagraph of each family switch places. Mutation meanwhile will swap gen between randomly selection of time
and fitness is calculated when both operations are functioning. Active chromosome will survive and passive chromosome will be tossed.

Colorni (1998) chosen to implement a set of activities that contains all information of a lecturer. These sets of activity represents timetable problems as a matrix which each column displays time within a week and rows indicates lecturer. Each element inside the matrix is gen and the differences value of allelic is based on sets of activities by each lecturers.
2.11 Strengths and Weaknesses of the Approaches

The table below shows the summary of the strengths and weakness of the approaches:

**Table 2.2: Summary of strengths and weaknesses of the approaches**

<table>
<thead>
<tr>
<th>Approaches</th>
<th>Strengths</th>
<th>Weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heuristic approach</td>
<td>The number of probability may be limited so the solution may be made within an appropriate time.</td>
<td>Subjects that could be allotted with time slots (no suitable time slot)</td>
</tr>
<tr>
<td>Integer Programming Approach</td>
<td>To aid in obtaining the most effective decision by utilizing all available resources</td>
<td>Too many or all variable in the formula is integer in value and normally the size of the problem is big.</td>
</tr>
<tr>
<td>Graph Coloring Approach</td>
<td>No conflict occurred because use difference coloring for every subject</td>
<td>Could not represent certain elements such as connecting two different subjects with different time.</td>
</tr>
<tr>
<td>Network Streaming Approach</td>
<td>Allocate time for the combination of lecturer-classroom whereby no conflicts generated in the schedule</td>
<td>Not suitable for analysis</td>
</tr>
<tr>
<td>Logical Constraints Arithmetic Approach</td>
<td>No timetable conflict</td>
<td>High constraints must be adhered to whilst some low constraints can be relaxed</td>
</tr>
<tr>
<td>Knowledge Base Approach</td>
<td>Priority is given to subject and selection of time slotting activity for allocating a reasonable time slot.</td>
<td>Every activities and time slots are represented separately</td>
</tr>
<tr>
<td>Tabu Searching</td>
<td>Minimizing the amount of conflicts</td>
<td>Obtaining high optimal solution</td>
</tr>
<tr>
<td>Annealing Simulation</td>
<td>Minimize the probability to transpire universal concentration</td>
<td>Involves low resolution compared to gradual cooling</td>
</tr>
<tr>
<td>Al-Khawarizmi Genetic</td>
<td>Able to solve many low constrictions, which almost result in an optimum solution. Makes selection using probability and does not use the rules as specified. Uses a coded parameter set. Evaluates objective functions using the population set</td>
<td>Not able to solve high constrictions problem</td>
</tr>
</tbody>
</table>
2.12 Current Research Of Other Researchers

One or two research studies published uses the Al-Khwarizmi Genetic as models to solve the university timetable problem. Both uses different problem representative.

2.12.1 Timetable Tuples

Burke et al (2000), represented each period as chromosome. Each chromosome contains genes which store all information on the timetable. Al-Khawarizmi Genetic in Figure 2.7 hereunder show the example of a timetable as a group of tuples. Each period will have several tuples known as label and the said tuple contains information on class, lecturer and rooms which has been ascertained.

<table>
<thead>
<tr>
<th>Label 1</th>
<th>Label 2</th>
<th>Label 3</th>
<th>Label 4</th>
<th>Label n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 4</td>
<td>Class 1</td>
<td>Class 9</td>
<td>Class 4</td>
<td>Class 3</td>
</tr>
<tr>
<td>Lecturer 5</td>
<td>Lecturer 2</td>
<td>Lecturer 5</td>
<td>Lecturer 9</td>
<td>Lecturer 3</td>
</tr>
<tr>
<td>Room 12</td>
<td>Room 4</td>
<td>Room 11</td>
<td>Room 8</td>
<td>Room 5</td>
</tr>
</tbody>
</table>

Figure 2.7: Timetable Using Tuples
Each label will be given a suitable period. The Al-Khawarizmi Genetic below shows that each period may contain as many tuples as possible.

\[\text{Period 1} \quad \text{Period 2} \quad \text{Period N}\]

\[
\begin{array}{c}
\text{Tuple} \\
10 & 67 \\
11 & 32 \\
\end{array} 
\quad 
\begin{array}{c}
\text{Tuple} \\
12 & 89 \\
62 & 7 \\
\end{array} 
\quad 
\begin{array}{c}
\text{Tuple} \\
12 & 89 \\
62 & 7 \\
\end{array}
\]

\[N = \text{Period Per Week}\]

**Figure 2.8: Allotment of Tuples to period**

\[\text{Period} \quad \text{Period 2} \quad \text{Period N}\]

\[
\begin{array}{c}
10 \\
67 \\
45 \\
32 \\
\end{array} 
\quad 
\begin{array}{c}
12 \\
89 \\
7 \\
\end{array} 
\quad 
\begin{array}{c}
78 \\
23 \\
99 \\
\end{array}
\]

**Figure 2.8: Tuples using Chromosome**
Al-Khawarizmi Genetic in Figure 2.8 shows the tuples being allotted to chromosome. Each period represent the chromosome. If there are 45 period it means that there are 45 chromosome or individualal in population. This method of representation will result in a huge problem. Firstly, it is difficult to represent subject that has various periods. Each subject is deemed to have only one period.

Secondly, when the cross-over and mutation process occurs, problems pertaining to changing of labels will occur, there will be subjects that will vanish or scheduled more than once. A child born or a new timetable will no longer have solutions that are correct or valid. When this occurs, the chromosome will have a high level of activity.

2.12.2 Timetable Matrix

Carter (1986), uses a set of Activity A for each lecturer that must be executed. Among them are:

- using the symbol 1,2,3,4,.....,0 for 10 classes where the subject is being taught.
- Using symbol D for rest time as an indicator for replacement lecturer.
- Symbol P for professional duties.
- Symbol S for classes where time has been duly fixed.
- Symbol . indicates that the lecturer need not work.
- Symbol --- indicates that the lecturer may go on leave.

As such set activity A = { 1,2,3,4,5,6,7,8,9,0,D,P,S,..-}
### Table 2.3: Example of Matrix That Represents Timetable

<table>
<thead>
<tr>
<th>Lecturer - Subject</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Network Technology</td>
<td>----</td>
<td>.441</td>
<td>2333.</td>
<td>...22</td>
<td>.1100</td>
<td>40D41</td>
</tr>
<tr>
<td>Object Oriented Techniques</td>
<td>----</td>
<td>1101.</td>
<td>..66</td>
<td>1D660</td>
<td>1D6…</td>
<td>06D00</td>
</tr>
<tr>
<td>Advanced Issues in Programming</td>
<td>55S5.</td>
<td>DSS..</td>
<td>5SS..</td>
<td>DSSS5</td>
<td>DSS..</td>
<td>----</td>
</tr>
<tr>
<td>Requirement Acquisition and Modeling</td>
<td>30.78</td>
<td>.8956</td>
<td>.7102</td>
<td>43.41</td>
<td>.2596</td>
<td>----</td>
</tr>
<tr>
<td>Software Matrices</td>
<td>0SSS.</td>
<td>SSSS0</td>
<td>1SSS0</td>
<td>----</td>
<td>5S55</td>
<td>11666</td>
</tr>
<tr>
<td>Software Reuse</td>
<td>D4242</td>
<td>332..</td>
<td>4D2..</td>
<td>----</td>
<td>3443.</td>
<td>342..</td>
</tr>
</tbody>
</table>

This activity represents problem in one matrix schedule \( R \) (\( m \times n \) where \( r_{ij} \in \mathbb{C}^{\square} \), each row indicates the lecturer and the row is time. Each element \( r_{ij} \) in matrix \( R \) is variable or gene. It may differ depending on subset \( A \) which indicates the lecturer at each row. From the Al-Khawarizmi Genetic herein, each lecturer will have a set of activity. For example:

- **Lecturer A = \{ ----..4412333…22.110040D41 \}**
- **Lecturer B = \{ ----1101….661D6601D6…06D00 \}**

This matrix must be feasible by fulfilling the constrictions stipulated. Each row will only have one class. For example, if on Friday, the first period for class 1 the subject taught is Advanced Network Technology, hence in the second period the subject will be Object Oriented Techniques. Class 1 should not have two subjects at any one time as it will indicate that there is more than one lecturer in class 1 at a time. The row representing period may be used to examine each class to ensure there is a subject, otherwise it will...
indicate that the said class has no lecturer and this is certainly not allowed in a feasible timetable.

Problem representation by Carter relates to only one lecturer who will be teaching one subject only. It differs from the teaching system in Malaysia which involves one subject that will be taught by several lecturers. For example, the subject Object Oriented Techniques, more than one lecturer conduct that subject to meet the many classes. In Malaysia there is also lecturer who teaches more than one subject.

2.13 Conclusion

This chapter depicts several intellectual researches technique regarding the scheduling difficulties. The primary objective in preparing timetable is obtaining conflict free for each activity sharing the same resources. Most of the method faces difficulties in generating a feasible timetable which faces long processing period without complying with several conditional constraints. This thesis program uses Al-Khawarizmi’s Genetic method to solve the timetable difficulties at the university.